Available online at www.sciencedirect.com
SCIENCE DIRECT®
International Journal of Approximate Reasoning
40 (2005) 147–180
INTERNATIONAL JOURNAL OF APPROXIMATE REASONING
www.elsevier.com/locate/ijar

An uncertainty interchange format with imprecise probabilities ☆

Pietro Baroni [a], Paolo Vicig [b,*]

[a] University of Brescia, Dip. Elettronica per l'Automazione, Via Branze 38, I-25123 Brescia, Italy
[b] University of Trieste, Dip. Matematica Applicata "B. de Finetti", Piazzale Europa 1, I-34127 Trieste, Italy

Received 1 September 2004; accepted 23 March 2005
Available online 18 April 2005

Abstract

This paper addresses the problem of exchanging uncertainty assessments in multi-agent systems. Since it is assumed that each agent might completely ignore the internal representation of its partners, a common interchange format is needed. We analyze the case of an interchange format defined by means of imprecise probabilities, pointing out the reasons of this choice. A core problem with the interchange format concerns transformations from imprecise probabilities into other formalisms (in particular, precise probabilities, possibilities, belief functions). We discuss this so far little investigated question, analyzing how previous proposals, mostly regarding special instances of imprecise probabilities, would fit into this problem. We then propose some general transformation procedures, which take also account of the fact that information can be partial, i.e. may concern an arbitrary (finite) set of events.

Keywords: Uncertainty transformations; Multi-agent systems; Imprecise probability theory; Pignistic probability; Partial possibilities

☆ This paper extends and generalizes the earlier results presented in [4,5,7].
* Corresponding author. Tel.: +39 040 5587108; fax: +39 040 54209.
E-mail addresses: baroni@ing.unibs.it (P. Baroni), paolo.vicig@econ.units.it (P. Vicig).

0888-613X/$ - see front matter
doi:10.1016/j.ijar.2005.03.001

1. Introduction

The definition of an interchange format for information exchange is a key issue in order to enable interoperability among independently developed and heterogeneous software systems, possibly adopting different internal representations. In particular, interoperability is a fundamental requirement in the development of *multi-agent systems* [30], namely systems composed by a set of autonomous, but interacting software entities (called *agents*). A typical application scenario is a virtual marketplace, where independently owned software agents automate some of the activities related to the buying and selling of goods [26].

The most influential proposal in this area is the knowledge interchange format (KIF) [24], which is a language designed for use in the interchange of knowledge among disparate computer systems. Actually, KIF is a prefix version of the language of first order predicate calculus with various extensions to enhance its expressiveness. As to our knowledge, KIF does not address the case where the pieces of information to be exchanged are fuzzy and/or affected by uncertainty. Similarly, a recently proposed tool to support interoperability among software agents [10] focuses on message translation into a common XML format, without considering the possible presence of uncertainty. This clearly restricts the applicability of this and similar proposals to the contexts where the importance of fuzziness and uncertainty is negligible.

To fill this gap, we address the problem of defining an interchange format for the exchange of uncertain information among heterogeneous agents, each one featuring a specific uncertainty representation.

Next to identifying a candidate interchange format, it is necessary to define transformation procedures between the common format and the agents' specific ones, in order to allow message exchanges. This issue turns out to be rather problematic, as generally valid transformation criteria cannot be identified, since, in particular, some reasonable requirements are indeed conflicting and, as a consequence, some arbitrariness is inherent to any transformation procedure.

The present work aims therefore at analyzing the main issues involved in the definition of an uncertainty interchange format and of the relevant uncertainty transformations, presenting a proposal based on imprecise probabilities and pointing out the relevant open problems.

The paper is organized as follows. Section 2 describes the problem of uncertainty interchange and the assumptions which are made to better specify it. Section 3 recalls some basic concepts and motivates the choice of coherent imprecise probabilities as an uncertainty interchange format, while in Section 4 the basic issues related to the definition of transformation procedures are discussed. Sections 5–7 discuss the problem of transforming imprecise probabilities into precise probabilities, possibilities, and belief functions, respectively. An example of application of these procedures is provided in Section 8, while Section 9 concludes the paper.

2. The problem of uncertainty interchange

We make the *basic assumption* that an agent might completely ignore the internal representation adopted by its partners and should not need to negotiate preliminarily any aspect of the interchange. Therefore no direct transformation between the representations internally adopted by different agents is possible and an intermediate interchange format is needed. While this assumption prevents the potential advantages of a tailored communication, it is coherent with the goal of maximum generality: for instance, it allows indirect communication through shared data bases, where the agent inserting new information does not know a priori who will eventually access it. Moreover, suppose the information exchange concerns n different uncertainty representations. Then direct transformations may require defining up to n^2–n transformation procedures, while at most $2n$ transformation procedures are needed if an interchange format is introduced.

Consider now the process of exchanging a piece of uncertain information between a sender, called agent S, and a receiver, called agent R. We assume that each agent X uses a specific internal representation language L_X, whereas the common interchange format language is denoted by L_{IF}. Agent S produces the information to be exchanged and transforms it from L_S into L_{IF}. The transformation result is then transmitted to agent R, which is in charge of transforming it from L_{IF} into L_R. Since L_{IF} should be, in general, more expressive than any specific language L_X, any transformation of the kind $L_X \to L_{\mathrm{IF}}$ should involve a null (or minimal) distortion or loss of information. On the other hand, any transformation $L_{\mathrm{IF}} \to L_X$ may involve some possibly significant information loss or distortion, since it goes from a general format to a more specific one.

To better specify the interchange problem, the following assumptions are made:

(1) the information exchange concerns uncertainty judgements[1] about (binary) non-conditional events;
(2) all agents share a common finite universe of discourse (or partition) Ω, and the events about which they formulate their uncertainty judgements constitute an arbitrary subset S of 2^{Ω}, the powerset of Ω.

Both assumptions are partially limiting and not completely realistic, but are useful to face gradually the problem; it is also important to note that, as will be stressed later, they are not forced by the interchange format we adopt.

A judgement qualifying the belief attitude of the agent with respect to the truth of an event (the judgement about its falsity can be referred to the truth of its negation) may be either *absolute* (when expressed through some quantification) or *relational* (also called comparative or qualitative, when defining a relation comparing the belief attitudes of two or more events). These two kinds of judgements cannot be easily

[1] This notion will be better specified in the sequel.

converted into each other, and in order to limit the scope and the extension of this paper, we focus here on absolute judgements.

In particular, we consider both precise and imprecise judgements. Imprecise judgements, i.e. judgements which associate a set of quantifications with an event, may be either crisp or fuzzy, depending on whether the sets of quantifications are, respectively, classical or fuzzy sets.

Clearly, fuzzy imprecise quantifications are more general and therefore appear to be a good basis for the definition of the interchange format. However, two difficulties have to be acknowledged:

- since a fuzzy set can in general be defined by any membership function, the formalism may turn out to be very complex, unless the class of admissible membership functions is constrained in some way;
- some of the most known uncertainty theories (imprecise probabilities, belief functions, possibility theory) can be characterized in terms of crisp imprecise quantifications, namely intervals.

For these reasons, while recognizing the potential importance of fuzzy sets for future developments, we limit the scope of the present paper to quantifications through crisp intervals, namely through a couple of numbers $(a,b) : a \leqslant b$; $a, b \in [0,1]$.

After deciding the format for a single uncertainty judgement, a standard format for an exchange of uncertain information should be defined. We assume that the agents adopt one of the existing standards for inter-agent communication, such as KQML [22] or FIPA ACL [23]: the choice of the communication language is rather indifferent, since our work focuses on the representation of the information carried by the message rather than on the structure of the message itself. Clearly these two aspects are independent.

As for the amount of information to be exchanged, S may be any set of events an agent considers interesting. This has, in particular, obvious advantages concerning the volume of information exchanged and the computational burden involved by the transformation. Moreover no non-requested information needs to be introduced in the exchange.

A message contents carrying uncertain information is therefore constituted by the specification of one or more events, which can be provided using KIF, and by the specification of the relevant uncertainty judgements, according to the interchange format that will be defined in the following.

3. Selecting an interchange format

3.1. Preliminaries

We use the symbol Ω to denote both the certain event and a *finite* set—also called *universal set* or *partition*—of pairwise disjoint (non-impossible) elementary events whose union is the certain event: $\Omega = \{\omega_1, \dots, \omega_N\}$.

Then $\omega_1, \ldots, \omega_N$ are called *atoms*, 2^Ω is the powerset of Ω, $|A|$ is the cardinality of $A \in 2^\Omega$, i.e. the number of distinct atoms of Ω whose union is A.

An *uncertainty assessment* (or *assignment*) is, in general, a function from an arbitrary set of events to the set of admissible uncertainty judgements. To recall the basic concepts underlying several well-known uncertainty theories it is however useful to initially consider the case where the set of events coincides with 2^Ω and the uncertainty assessment is a function $f: 2^\Omega \to [0,1]$.

An uncertainty assessment should at least satisfy a minimal monotonicity requirement and two obvious conditions: this gives rise to the notion of capacity.

A mapping $C: 2^\Omega \to [0,1]$ is a (normalized) *capacity* [9] whenever:

$$C(\emptyset) = 0; \quad C(\Omega) = 1; \quad C(A) \leqslant C(B), \quad \forall A, B \in 2^\Omega \text{ such that } A \subset B.$$

A capacity is 2-*monotone* iff $\forall A, B \in 2^\Omega,\ C(A \cup B) \geqslant C(A) + C(B) - C(A \cap B)$.

Belief functions [41], and in particular necessity measures [19] and, when defined on 2^Ω, finitely additive probabilities, are special cases of 2-monotone capacities. Their definitions are well known; we shall recall in Proposition 1 their characterizations in terms of their Möbius inverses [9].

For any $f: 2^\Omega \to \mathbb{R}$, there is a one-to-one correspondence between f and its *Möbius inverse* or *mass function* $m: 2^\Omega \to \mathbb{R}$, given $\forall A \in 2^\Omega$ by [9,40]:

$$m(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} f(B), \quad f(A) = \sum_{B \subseteq A} m(B) \tag{1}$$

The events $A \in 2^\Omega$ such that $m(A) \neq 0$ are called *focal elements*.

Proposition 1. *Given $f: 2^\Omega \to \mathbb{R}$, let m be its Möbius inverse. Then*

(a) *f is a* capacity *iff m is such that*: $m(\emptyset) = 0$; $\sum_{B \in 2^\Omega} m(B) = 1$; $\forall A \in 2^\Omega, \forall \omega \in A$, $\sum_{\omega \in B \subseteq A} m(B) \geqslant 0$.
Further, if f is a capacity *and F the set of its focal elements, then*
(b) *f is a* 2-monotone capacity *iff* $\forall A, B \in 2^\Omega,\ \sum_{C \subset A \cup B, C \nsubseteq A, C \nsubseteq B} m(C) \geqslant 0$;
(c) *f is a* belief function *iff m is non-negative*;
(d) *f is a* necessity measure *iff F is totally ordered by relation* '$\subset$';
(e) *f is a* (*precise*) probability *iff* $(A \in F) \Rightarrow (A$ *atom of* $\Omega)$.

Note that if f is a capacity, $f(\omega) = m(\omega) \geqslant 0,\ \forall \omega \in \Omega$.

The *conjugate* C' of a capacity C is defined by the conjugacy relation $C'(A) = 1 - C(A^c),\ \forall A \in 2^\Omega$. The conjugate of a 2-monotone capacity, a belief function (*Bel*), a necessity measure (N) is termed, respectively, 2-alternating capacity, plausibility function (*Pl*), possibility (Π). A precise probability (P) is self-conjugate.

Capacities may be too loose as uncertainty measures, as stronger consistency requirements are often needed. Coherent lower (and upper) imprecise probabilities are a natural candidate, since they are more general than the measures, other than capacities, recalled so far and their definition is based on well-founded rationality requirements. The notion of *coherent lower probability*[2] is defined in [47, Section 2.5], referring to an *arbitrary* (finite or not, structured or not) set of events S. We assume here that $S \subseteq 2^{\Omega}$. Coherent lower probabilities are indirectly characterized as lower envelopes of precise probabilities on S. Denoting with $\mathscr{M}$ the set of all precise probabilities *dominating* $\underline{P}$ on S (i.e. $P \in \mathscr{M}$ iff $P(A) \geqslant \underline{P}(A),\ \forall A \in S$), the following *lower envelope theorem* holds [47].

Proposition 2. *$\underline{P}$ is a coherent lower probability on S iff there exists a (non-empty) set $D \subseteq \mathscr{M}$ such that $\underline{P}(A) = inf_{P \in D} P(A),\ \forall A \in S$ (inf is attained).*

One may refer to either lower ($\underline{P}$) or upper ($\overline{P}$) probabilities only, exploiting the conjugacy relation:

$$\underline{P}(A) = 1 - \overline{P}(A^c) \tag{2}$$

In particular, assessing both $\underline{P}(A)$ and $\overline{P}(A)$ is equivalent to assessing $\underline{P}(A)$ and $\underline{P}(A^c) = 1 - \overline{P}(A)$.

When $\underline{P}(A) = \overline{P}(A) = P(A),\ \forall A \in S,\ P$ is a precise probability (coherent by de Finetti's definition [15]) on S (a finitely additive probability if $S = 2^{\Omega}$).

Later, we shall use the following necessary condition for coherence:

$$\underline{P}(A \cup B) \geqslant \underline{P}(A) + \underline{P}(B), \quad \forall A, B : A \cap B = \emptyset \quad (A, B, A \cup B \in S) \tag{3}$$

When $S = 2^{\Omega}$, lower (upper) probabilities are capacities, and include 2-monotone (2-alternating) capacities as special cases.

Coherent imprecise probabilities can therefore be regarded as a very general tool, which generalizes various uncertainty measures. In particular, a belief function [41] is a special case of lower probability [48], and so is a necessity measure which can be seen as a special case of belief function (actually, as a consonant belief function) [19,42], while plausibility and possibility measures are special upper probabilities [48].

Imprecise probabilities appear to be a suitable candidate for the definition of an uncertainty interchange format, for the following reasons.

Firstly, no transformation is needed from the agent internal representation L_X to the common interchange format whenever L_X is based on an uncertainty measure which is a special case of imprecise probability, like the (quite common) ones we consider in this paper. In fact, in such instances, information produced by the agent may be simply read as an imprecise probability in the interchange format, without modifying any of its numerical values.

Moreover, some of the assumptions we made might be widely relaxed while keeping on using an interchange format based on imprecise probabilities. In fact, coherent lower (and upper) probabilities are defined also on infinite sets of events. Further, by the extension theorem [47], an imprecise probability on any set of events S can always be coherently extended to any superset of S, and this allows exchanging infor-

[2] We shall often omit the term coherent in the sequel.

mation in a dynamic setting where the universe of discourse is not fixed. These features are shared also by generalizations of imprecise probabilities to conditional imprecise probabilities and previsions (the latter are suited for handling information on conditional random variables) [47].

Of course, there are important ways of expressing uncertainty which are not special cases of imprecise probabilities and therefore would need some transformation before using the proposed interchange format: we mention fuzzy judgements, which should be reduced to crisp intervals, and comparative probabilities, for which a realization problem (by means of an imprecise probability) arises, and it is not guaranteed a priori that it always has a solution. These issues are beyond the scope of the present paper.

Another question, which can be partly encompassed within the theory of coherent imprecise probabilities, concerns uncertainty assessments which are themselves imprecisely determined. For instance, interval-valued belief structures were introduced in [17] to extend the theory of belief functions to the case of beliefs expressed with some imprecision. Suppose that a lower uncertainty assessment on S is imprecisely specified, meaning that there is a set F_L of assessments f which are all admissible. This type of uncertainty may be described computing $f_L^{(l)}(A) = \inf_{f\in F_L} f(A)$ and $f_L^{(u)}(A) = \sup_{f\in F_L} f(A),\ \forall A \in S$. It is known that whenever the measures f are coherent lower probabilities, so is their lower envelope $f_L^{(l)}$ (cf. [47, Theorem 2.6.3 (b)]); in particular $f_L^{(l)}$ is a coherent lower probability when all $f \in F_L$ are (possibly partial, as will be the case later on) belief functions or necessities. On the contrary, function $f_L^{(u)}$ is generally not a coherent imprecise probability: hence imprecise probabilities can describe this kind of imprecision partly, but generally not fully (the description is complete when all measures f are precise probabilities). Specular results obviously apply to the lower and upper bounds on any set of upper uncertainty measures.

4. Defining uncertainty transformations

4.1. Transformation criteria

The problem of transforming an uncertainty assessment U_{ORIG} expressed in the context of a given formalism L_{ORIG} into another assessment U_{DEST} within a different formalism L_{DEST} has no straightforward solution. In fact:

- if L_{DEST} is intrinsically more precise than L_{ORIG} (for instance, L_{DEST} is a precise probability while L_{ORIG} is an imprecise one), the precision imposed by L_{DEST} has to be achieved by introducing some constraint, not entailed by the original representation;
- if L_{DEST} is intrinsically less precise than L_{ORIG} (as in the case, to be discussed later, where L_{DEST} is a possibility while L_{ORIG} is an imprecise probability), the precision inherent to the original representation needs to be sacrificed in some way.

In both cases, some degree of arbitrariness is unavoidable and any transformation mechanism turns out to be questionable in some respect (see, for instance, the discussion in [21] about precise probability/possibility transformations). As a consequence, several different transformation criteria have been considered in the literature. They can be roughly partitioned into two main classes: *acceptability criteria* define necessary conditions to be satisfied by acceptable transformation results, while *selection criteria* can be used to select a transformation result among the acceptable ones.

Examples of acceptability criteria are the consistency criterion, the preference preservation principle, the uncertainty invariance principle.

The *consistency criterion* requires that U_{DEST} respects some consistency constraints, referring to U_{ORIG}. For instance in [21] consistency between a precise probability P and a possibility Π is represented by the following dominance condition: $P(A) \leqslant \Pi(A),\ \forall A \in 2^{\Omega}$. An analogous condition is adopted in [29] to define consistency between a precise probability and a capacity and in [20] in the context of approximations of belief functions.

More generally, when considering two uncertainty formalisms L_1 and L_2 it may appear that L_1 is intrinsically more precise than (or at least as precise as) L_2 (e.g. a precise probability is clearly more precise than a possibility). In this case, a transformation from L_1 into L_2 should give a more imprecise evaluation (vice versa when passing from L_2 to L_1). This leads operationally to a dominance condition which can be expressed in terms of upper measures:

$$\overline{L}_1(A) \leqslant \overline{L}_2(A), \quad \forall A \in 2^{\Omega} \tag{4}$$

or equivalently in terms of their conjugate lower measures:

$$\underline{L}_1(A) \geqslant \underline{L}_2(A), \quad \forall A \in 2^{\Omega} \tag{5}$$

Preference preservation requires that some credibility order induced on events by U_{ORIG} is preserved by U_{DEST}. An example is the strict preference preservation principle, considered in [21]: $U_{\mathrm{ORIG}}(A) > U_{\mathrm{ORIG}}(B) \Longleftrightarrow U_{\mathrm{DEST}}(A) > U_{\mathrm{DEST}}(B)$, $\forall A, B \in 2^{\Omega}$.

In general, however, associating a (partial) credibility ordering with an uncertainty assessment may be itself problematic, as discussed in [46] for the case of belief functions, and gives operationally rise to a large number of constraints.

The *uncertainty invariance* principle, proposed in [33], states that uncertainty transformations should not modify the information contents of a given assignment. To apply this principle, a measure of information has to be defined for the uncertainty measures involved in the transformation process.

Turning to selection criteria, the quantitative similarity criterion and the specificity preservation criterion can be considered.

The *quantitative similarity* criterion is an intuitive requirement, imposing the minimization of some distance between U_{ORIG} and U_{DEST}.

Specificity preservation is somewhat related to uncertainty invariance: the underlying idea is that in a transformation from a less precise formalism to a more precise one, the amount of "precision" arbitrarily added (i.e. specific information not entailed by U_{ORIG}) should be minimized, as, conversely, the amount of additional imprecision should be minimized in the opposite transformation. For instance maximum entropy is a well-known criterion to select a minimally specific precise probability [29], while maximal specificity has been required for transformations from probabilities into possibilities [21].

Though each of the above listed criteria has an intuitively sound justification, two main problems are worth pointing out:

- not all criteria are currently applicable to any transformation: as already noticed above, the notions underlying some criteria (e.g. information measure, entropy, specificity) may be undefined in some theories;
- some criteria are inherently conflicting, for instance the constraints imposed by preference preservation may be very strong and tend to significantly widen the imprecision gap introduced in a transformation, thus contrasting with other criteria such as uncertainty invariance, quantitative similarity and specificity preservation.

The reader may also refer to [12] for a survey and analysis of various procedures and of the relevant criteria, in the case of transformations from belief functions into precise probabilities. In particular it is remarked that no existing transformation procedure complies with all the criteria analyzed. For this reason, no transformation procedure appears to be *optimal* in an absolute sense, since it can be rated differently, depending on the relative importance one ascribes to conflicting criteria.

The criteria adopted for the transformation procedures we propose will be discussed in the relevant sections.

4.2. Coping with theory peculiarities

Other problematic aspects which may affect transformation definition and behavior are related with peculiar and somewhat "incompatible" features of different theories.

For instance, basic properties of possibility theory lead to discontinuities which are not present in other formalisms. As a simple example consider the problem of transforming the lower and upper probability of a single event A (say a probability interval: $[\underline{P}(A), \overline{P}(A)]$) into a necessity and a possibility value: $[N(A), \Pi(A)]$. As well known, $[N(A), \Pi(A)]$ is constrained to have the form $[0, \Pi]$ or $[N, 1]$. This means that $N(A), \Pi(A)$, viewed as an imprecise probability, is always either lower (if $N = 0$) or upper (if $\Pi = 1$) maximally imprecise. As a consequence, a sort of stretching of the probability interval in one direction is required for the transformation. This can be obtained by transforming $\overline{P}(A)$ into a possibility value 1, if $1 - \overline{P}(A) < \underline{P}(A)$, i.e. if $\overline{P}(A)$ is closer to 1 than $\underline{P}(A)$ is to 0, or transforming $\underline{P}(A)$ into a necessity value 0, if $1 - \overline{P}(A) > \underline{P}(A)$. The case

$$1 - \overline{P}(A) = \underline{P}(A) \tag{6}$$

shows however a singularity of this method, that is present, when $P(A) = 0.5$ (the special case of (6) with $\overline{P}(A) = \underline{P}(A)$), also in the transformation proposed for precise probabilities in [21]. In fact, uniform probability is equated, in possibility theory, to *total ignorance*: in absence of any preference, we get the extreme assignment $\Pi(A) = 1, N(A) = 0$ (and hence $\Pi(A^c) = 1, N(A^c) = 0$), which should also be the transformation of an imprecise probability assignment obeying (6), since again we have no reason for modifying the upper rather than the lower imprecise probability. However in these cases the probability/possibility transformation operator behaves discontinuously. To exemplify, put $\underline{P}(A) = 0.5,\ \overline{P}(A) = 0.5 + \varepsilon$, with a quite small positive ε (being therefore close to the total ignorance case): this gives $N(A) = 0.5 + \varepsilon,\ \Pi(A) = 1$ (and hence $N(A^c) = 0,\ \Pi(A^c) = 0.5 - \varepsilon$) with a large discontinuity of $N(A)$ (and of $\Pi(A^c)$) as $\varepsilon \to 0+$. This shows that even in the simplest case of a single event, uncertainty transformations between different measures may involve some singular behaviors and unavoidable distortions.

As another example, consider the role of the Möbius inverse (or mass function): in several theories its non-negativity is regarded as a basic property, so that mass values can be interpreted as "degrees of evidence". Examples of information measures [33] and transformation procedures [20,43] based on this property are available in the literature. However, these proposals have no counterpart in the context of imprecise probabilities, where masses can be negative and have no clear intuitive interpretation [6].

As a third example, it may be the case with imprecise probabilities that:

$$\overline{P}(A) < \overline{P}(B) \wedge \overline{P}(A \cup C) > \overline{P}(B \cup C), \quad A \cap C = B \cap C = \emptyset,$$

while these conditions cannot hold with precise probabilities. These *preference inversions* may affect the ability of a given transformation to minimize additional imprecision, as discussed in [4].

4.3. Partial assessments

As stated in Section 2, we do not require that the information exchange concerns a complete uncertainty assignment defined on 2^{Ω}, but rather a "partial" assessment on an arbitrary set of events $S \subset 2^{\Omega}$. This poses two main problems.

First of all, most uncertainty theories are customarily defined referring to (at least) an algebra of events. In order to enable partial exchanges it is therefore necessary to characterize partial uncertainty assignments within each of the considered theories: this issue will be dealt with in the relevant sections.

A more subtle problem is related to the implicit information possibly carried by a partial assessment: one may wonder what does a given partial assessment on S entail on other events, not belonging to S.

Concerning this problem, it is known that coherent lower probabilities can *always* be coherently extended on any $S' \supset S$. In particular, when considering one additional event A, the set of all coherent extensions of $\underline{P}$ to A is a closed (non-empty) interval $[\underline{P}_E(A), \underline{P}_U(A)]$, where $\underline{P}_E(A)$ is the *natural extension* [47] of $\underline{P}$ to A. That is $\underline{P}_E(A)$ is the *least committal* or vaguest admissible coherent extension of $\underline{P}$ on A. It may be obtained as the infimum value on A of all precise probabilities dominating $\underline{P}$ on S, which requires solving a linear programming (LP) problem. In principle, other coherent extensions may be interesting, especially the *upper extension* $\underline{P}_U(A)$ which has the opposite meaning of least vague coherent extension of $\underline{P}$. However, as shown in [45, Section 5], computing $\underline{P}_U(A)$ may require solving up to $|S|$ distinct LP problems. The natural extension has also another advantage: if we compute separately $\underline{P}_E(A_i)$ for $i = 1, \dots, m$ (i.e. we find separately the natural extension of $\underline{P}$ on $S_i = S \cup \{A_i\}$), then $\{\underline{P}_E(A_1), \dots, \underline{P}_E(A_m)\}$ is the natural extension of $\underline{P}$ on $S \cup \{A_1, \dots, A_m\}$. This is generally not true for any other coherent extension: for instance, if we compute $\underline{P}_U(A_1)$ and put $\underline{P}(A_1) = \underline{P}_U(A_1)$, then the upper extension $\underline{P}^*_U(A_2)$ of $\underline{P}$ from $S \cup \{A_1\}$ to $S \cup \{A_1, A_2\}$ will be generally less than the value $\underline{P}_U(A_2)$ representing the upper extension of $\underline{P}$ from S to $S \cup \{A_2\}$.

It can then be questioned whether one should apply a transformation procedure directly on a given partial assessment or should first compute its natural extension (on 2^{Ω}) and then transform it.

The direct way appears more appealing at a first glance. However, the choice between direct and indirect way is in general not straightforward.

On one hand, the direct way avoids or reduces the computations required to determine the natural extension. On the other hand, a partial assignment contains some implicit information which is actually ignored by the direct way, but could affect transformation results if considered. Therefore ignoring *all* implications of a given partial assessment might be expected to originate an unsatisfactory transformation result. Again, it does not seem that a univocal answer to this question can be easily given. An example will be discussed in Section 5.2. Note also that when a partial assessment is an imprecise probability with specific properties (e.g. it is a partial possibility) it is not guaranteed, in general, that such properties are preserved by its natural extension.

In the following sections we deal with the problem of defining a transformation procedure from an imprecise probability into a precise probability, a possibility and a belief function respectively. For each transformation we will:

- analyze transformation procedures previously proposed in the literature;
- consider the partial assessment case;
- define a transformation procedure which can be applied to both complete and partial assessments.

5. Transformations into precise probabilities

5.1. Previous proposals

The problem of transforming an imprecise probability into a precise probability has not been previously considered in the literature in its full generality. Several proposals exist however for transformations from specific subclasses of imprecise probabilities. In the following we briefly discuss them, analyzing in particular whether they can be generalized from the context where they were initially conceived.

5.1.1. Voorbraak's Bayesian transformation

Voorbraak's Bayesian transformation takes as input the mass function m of a given belief function and outputs a mass function m_V which, in [46], is defined to be zero for all non-atomic events of 2^Ω, while for every atom $\omega \in \Omega$ it is:

$$m_V(\omega) = \frac{\sum_{A \supseteq \omega} m(A)}{\sum_{B \subseteq \Omega} m(B) \cdot |B|} \tag{7}$$

Such a mass assignment corresponds to a precise probability $P_V(\omega) = m_V(\omega)$, $\forall \omega \in \Omega$ (cf. Proposition 1(e)). We note that this still holds if m is the mass function of a generic capacity, since again m_V is non-negative, as is not difficult to verify using Proposition 1(a). The transformation can be equivalently rewritten in terms of the initial plausibility values of atoms: $P_V(\omega) = m_V(\omega) = Pl(\omega)/\sum_{\omega \in \Omega} Pl(\omega)$.

In other words, Voorbraak's proposal is a normalization of the plausibility values of atoms of Ω (more generally, when applied to a capacity C, it normalizes the values on the atoms of its conjugate C'). As such, it clearly respects a preference preservation criterion *restricted* to the plausibility values of the atoms of Ω. This is a relatively weak property, in particular the transformation does not match with the consistency criterion, as pointed out e.g. in [20]. However this transformation shows appreciable properties in relation with Dempster's rule, as discussed in [11].

In [28] a method for approximating belief functions based on the concept of fuzzy T-preorder is proposed. However this method is mainly focused on approximating a belief function by means of possibility distributions, even though it could be directly applied to transforming coherent upper probabilities. The existence and uniqueness of a probability which approximates the belief function is guaranteed only for a specific choice of the T-norm to be used in the transformation. In this case, as shown in [28], the probability obtained coincides with Voorbraak's, therefore the same considerations can be applied.

5.1.2. Pignistic probability

The pignistic probability transformation (PPT) has been considered in a variety of papers, mainly concerning belief functions (e.g. [20,43]).

In particular, pignistic probabilities are suitable for decision making under expected utility theory in the transferable belief model (TBM). The necessity of PPT in this context is justified by a linearity requirement (for recent results on this see [44], where the formal analogy between the pignistic probability and the Shapley value in game theory is also highlighted).

The PPT takes in input a mass function m and produces a probability P_{pign} on the atoms of Ω:

$$P_{\text{pign}}(\omega) = m_{\text{pign}}(\omega) = \sum_{\Omega \supseteq A \ni \omega} \frac{m(A)}{|A|}, \quad \forall \omega \in \Omega \tag{8}$$

This transformation is based on the principle of insufficient reason applied to the focal events: it distributes uniformly their masses over their atoms.

Assuming now that m is the Möbius inverse of a capacity C, it is easy to see (using (1), (8) and additivity of P_{pign}) that

$$P_{\text{pign}}(B) = C(B) + \sum_{A \not\subseteq B} \frac{|A \cap B|}{|A|} m(A), \quad \forall B \in 2^{\Omega} \tag{9}$$

It is clear from (9) and Proposition 1(c) that PPT preserves the consistency criterion when C is a belief function. Now the point is: what if C is not a belief function? It is not even immediate that P_{pign} is then a probability, since negative masses appear in (8). The question was first addressed in [43], where P_{pign} was also derived from axioms for combining *credibility functions*, i.e. capacities with some additional axiomatical properties. We extend this result by proving in Proposition 4 that PPT returns a precise probability from any capacity. Preliminarily we summarize some results stated in [9].

Proposition 3. *Given a capacity $C : 2^{\Omega} \to \mathbb{R}$, an associated set of precise probabilities V can be defined as follows. Let Σ be the set of all permutations of the atoms of Ω, with $|\Omega| = N$. Given $\sigma \in \Sigma$, $\sigma = (\omega_{i_1}, \ldots, \omega_{i_N})$, define $\sigma_0 = \emptyset$ and, for $n = 1, \ldots, N$, $\sigma_n = \omega_{i_1} \cup \cdots \cup \omega_{i_n}$, $P_{\sigma}(\omega_{i_n}) = C(\sigma_n) - C(\sigma_{n-1})$. Let $V = \{P_{\sigma} : \sigma \in \Sigma\}$. Then*

(a) *Every P_{σ} is a (precise) probability and, $\forall \omega \in \Omega$,*

$$P_{\sigma}(\omega) = \sum_{\Omega \supseteq A \ni \omega} \lambda(A, \omega) m(A) \tag{10}$$

where $\lambda(A, \omega_{\sigma}(A)) = 1$, $\omega_{\sigma}(A)$ being the last element of A in permutation σ, and $\lambda(A, \omega) = 0$, $\forall \omega \neq \omega_{\sigma}(A)$.

(b) *The set V coincides with the set of the vertices of the set $\mathcal{M}$ of all precise probabilities dominating C if and only if C is 2-monotone.*

Proposition 4. *Let m be the mass function corresponding to a capacity C. Then P_{pign} as defined in (8) is a precise probability.*

Proof. Using (10), we have that $P_{\sigma}(\omega)$, and consequently $\sum_{\sigma \in \Sigma} P_{\sigma}(\omega)$, are a weighted sum of the masses of events A such that $A \ni \omega$. The weight of a generic $m(A)$ in $\sum_{\sigma \in \Sigma} P_{\sigma}(\omega)$ is equal to the cardinality of the set $\{\sigma' | \sigma' \in \Sigma, \omega = \omega_{\sigma'}(A)\}$, i.e. to the number of times $\lambda(A, \omega) = 1$. The cardinality of this set is actually $\frac{N!}{|A|}$, since any $\omega \in A$ has the same chance of being the last element of A in a given permutation σ. Therefore

$$\frac{1}{N!} \sum_{\sigma \in \Sigma} P_{\sigma}(\omega) = \frac{1}{N!} \sum_{\Omega \supseteq A \ni \omega} \frac{N!}{|A|} m(A) = P_{\text{pign}}(\omega), \quad \forall \omega \in \Omega \tag{11}$$

From (11) and Proposition 3(a), P_{pign} is a convex combination of precise probabilities, and is therefore itself a precise probability. $\square$

We answer now a further question concerning how PPT relates to the consistency criterion.

Proposition 5. *Let $C : 2^{\Omega} \to \mathbb{R}$ be a given capacity with mass function m, and P_{pign} be defined by (8).*

(a) *If C is 2-monotone, $P_{\mathrm{pign}}(A) \geqslant C(A),\ \forall A \in 2^{\Omega}$.*
(b) *If C is not 2-monotone, P_{pign} may or may not dominate C. If $\underline{P}$ is a coherent lower probability, P_{pign} dominates $\underline{P}$ on the atoms of Ω, i.e. $P_{\mathrm{pign}}(\omega) \geqslant \underline{P}(\omega),\ \forall \omega \in \Omega$, but not necessarily elsewhere.*

Proof. To prove (a), observe that if C is 2-monotone the probabilities P_{σ} are vertices of $\mathscr{M}$ by Proposition 3(b), and as such dominate C; then also their convex combination P_{pign} dominates C (see (11)).

If C is not 2-monotone, examples may be found where P_{pign} does not dominate C. Consider for instance the coherent lower probability $\underline{P}$ on 2^{Ω}, with $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}$, which is the lower envelope of the three precise probabilities P_1, P_2, P_3, determined by orderly assigning the following values on the atoms ω_1, ω_2, ω_3, ω_4, ω_5: P_1-values $[0.49, 0.35, 0.12, 0.01, 0.03]$, P_2-values $[0.14, 0.03, 0.07, 0.36, 0.40]$, P_3-values $[0.36, 0.05, 0.29, 0.14, 0.16]$. Then P_{pign} is given by the following values on the atoms ω_1, ω_2, ω_3, ω_4, ω_5: $[0.3198\overline{3}, 0.1631\overline{6}, 0.142\overline{3}, 0.172\overline{3}, 0.202\overline{3}]$. We have that $\underline{P}(\omega_1 \cup \omega_4) = 0.5 > P_{\mathrm{pign}}(\omega_1 \cup \omega_4) = 0.4921\overline{6}$.

To prove the second part of (b), let $\omega \in \Omega$. Considering a permutation σ (Proposition 3), let r be the position of ω in σ, i.e. $\omega_{i_r} = \omega$. Then $P_{\sigma}(\omega) = P_{\sigma}(\omega_{i_r}) = \underline{P}(\omega_{i_1} \cup \cdots \cup \omega_{i_r}) - \underline{P}(\omega_{i_1} \cup \cdots \cup \omega_{i_{r-1}}) \geqslant \underline{P}(\omega_{i_r}) = \underline{P}(\omega)$, using (3) in the inequality. Since σ is arbitrary, $P_{\sigma}(\omega) \geqslant \underline{P}(\omega),\ \forall \sigma$. Hence also the convex combination P_{pign} of the probabilities P_{σ} is such that $P_{\mathrm{pign}}(\omega) \geqslant \underline{P}(\omega)$. $\square$

We may conclude from Proposition 5 that PPT preserves the consistency criterion as far as it is applied to 2-monotone capacities. It may not preserve it outside 2-monotonicity, even though consistency may at least partially hold, as demonstrated in (b).

To get some empirical insight of the behaviour of PPT outside 2-monotonicity, we randomly generated a large number of coherent lower probabilities which were not 2-monotone (see [7] for details), computing also their corresponding P_{pign}: it has to be reported that the percentage of non-dominating P_{pign} was relatively low, and dominance violation numerically rather small. One of the examples we found is used in the proof of Proposition 5.

The pignistic probability has also been interpreted (e.g. in [21], which focuses on possibility measures) as center of gravity of the vertices of $\mathscr{M}$. This interpretation is clearly supported by Proposition 3(b) and (11), from which it is also patent that its validity is limited to 2-monotone probabilities. However the interpretation seems debatable even in the context of 2-monotonicity. In fact, from (11), P_{pign} is the average of $N!$ probabilities P_σ, which are *not* necessarily all distinct: distinct permutations in Proposition 3 may well originate the same probability (examples are easily found). This means that P_{pign} is actually a *weighted* average of the *distinct* P_σ, and the weight of any P_σ is given by the number of distinct permutations σ which give rise to it.

A point which seems therefore difficult to justify in the center of gravity interpretation is the meaning of the weights in terms of the initially assigned 2-monotone capacity.

5.1.3. Uncertainty invariant transformations

In order to apply the uncertainty invariance principle to imprecise probabilities, the definition of the *aggregate uncertainty* measure (AU), namely the maximum value of the Shannon entropy among all probability distributions dominating a given belief function [33], could be directly extended to coherent imprecise probabilities. Some limitations of AU are pointed out in [34]. In particular, AU does not distinguish among all imprecise probabilities which are consistent with the uniform probability, including the case of total ignorance.

A transformation based on the uncertainty invariance principle and using the AU measure consists in determining the maximum entropy precise probability among those in $\mathscr{M}$. This is equivalent to adopting the consistency criterion with maximum entropy as a selection criterion, but is intrinsically in contrast with preference preservation, tending to equate all probability values of atoms, as far as allowed by consistency. In our opinion, this is a drawback of the criterion, as for transformations devoted to uncertainty interchange. A discussion of pros and cons of maximum entropy methods may be found in [47, Section 5.12], where it appears that these methods may be appropriate in certain specific decision problems.

Apart from theoretical issues, as to our knowledge no algorithm has been devised for computing the maximum entropy probability P_{ME} consistent with an imprecise probability assessment. In the case of 2-monotone capacities, it is shown in [29] that P_{ME} is unique and an algorithm for computing it is provided. We checked that this algorithm does not ensure the consistency condition outside 2-monotonicity. Consider for this the lower probability assignment $\underline{P}$ in the proof of Proposition 5(b). The algorithm proposed in [29] produces, when applied to $\underline{P}$, a probability P_J determined by the values $[0.26, 0.16, 0.16, 0.16, 0.26]$ on the atoms of Ω. However P_J is not consistent with $\underline{P}$, since $P_J(a \cup d) = 0.42 < \underline{P}(a \cup d) = 0.5$.

5.2. An imprecise to precise probability transformation

As shown in the previous section, some existing transformation procedures preserve their applicability or some important properties only in the context of 2-monotone capacities. While 2-monotonicity arises in certain contexts, for instance when using pari-mutuel models (at racetracks or in life insurance) [47], or, more generally, convex functions of precise probabilities [16], it is also known that there are important situations which cannot be adequately described by 2-monotonicity [47, Section 5.13.4]. Moreover, an empirical analysis carried out in [7] indicates that 2-monotone probabilities cannot be considered an adequate representative of imprecise probabilities ("most" imprecise probabilities are not 2-monotone).

Another point from Section 5.1 is that known transformations often make use of the mass function m. This might suggest seeking for a transformation based on m in our context too. However the interpretation of m for imprecise probabilities is unclear (see [6] for a discussion). Moreover we are aware of no characterization of imprecise probabilities in terms of m (like those in Proposition 1).

Further, we shall be interested in transforming assessments on a generic (finite) set of events S,[3] while existing proposals refer to complete assignments. In particular, although the mass function exists also in the partial case [40], it is easy to see that it does not preserve the properties it has in a complete assignment.

We recall that the notion of coherent precise probability is well-established, and is defined on arbitrary sets of events [15].

We shall now illustrate a transformation procedure operating on any coherent imprecise probability defined on a set of events S, i.e. the events considered interesting by the agents involved in the interchange.

Defining $S^* = 2^{\Omega} \setminus \{\emptyset, \Omega\}$, we suppose at first $S = S^*$: the case $S \subsetneqq S^*$ will be considered later. We require the transformation to meet the consistency principle, which, in terms of upper and lower probabilities, imposes for the resulting precise probability P^* that

$$\underline{P}(A) \leqslant P^*(A) \leqslant \overline{P}(A), \quad \forall A \in S^* \tag{12}$$

When $|\Omega| = 2$ (hence $S^* = \{A, A^c\}$), P^* may be fully determined from $P^*(A) = \frac{\overline{P}(A)+\underline{P}(A)}{2}$. This seems reasonable, since P^* reduces then the imprecision of both $\underline{P}$ and $\overline{P}$ by the same amount, and there is no reason for P^* to be closer to either of $\underline{P}$ or $\overline{P}$. A straightforward generalization for $|\Omega| > 2$ of the idea of eliminating imprecision in a symmetric way for each event in S^* leads to considering $P_m = \frac{\overline{P}(A)+\underline{P}(A)}{2}$, $\forall A \in S^*$.

In general P_m is not a precise probability, but we may choose a probability P^* close to it in some way. Obviously there are several approximation choices; selecting in particular the common *least-squares* approximation of P_m leads to the following *transformation problem* (TP):

$$\min \varphi = \sum_{A \in S^*} \left(\sum_{\omega_i \in A} P^*(\omega_i) - P_m(A) \right)^2 \tag{13}$$

with the constraints (12) and

$$P^*(\omega_i) \geqslant 0, \ \ i = 1, \ldots, N; \quad \sum_{i=1}^{N} P^*(\omega_i) = 1 \tag{14}$$

[3] We may always assume that $\emptyset$ and Ω do not belong to S, since their uncertainty evaluation is trivial.

The variables in TP are $P^*(\omega_1), \ldots, P^*(\omega_N)$. Since φ is convex on $\mathbb{R}^{|\Omega|}$ and the set of constraints S_C is a (non-empty[4]) polyhedral set, TP is a convex quadratic programming problem, for which polynomial-time solving algorithms are known (see e.g. [8, Section 11.2]). TP has some desirable properties, which we derived using well-known results in calculus and convex programming:

(a) Problem TP always returns a unique P^*. In particular, TP detects whether P_m is a precise probability, since in such a case it gives $P^* = P_m$.
(b) It may be useful to solve the linear system which equates to zero the gradient vector of φ. In fact, if its (unique) solution is an interior point in S_C then it is the required P^*; otherwise we get to know that P^* will be equal to either $\overline{P}(A)$ or $\underline{P}(A)$ for at least one $A \in S^*$.
(c) If $\overline{P}(A) = 1,\ \underline{P}(A) = 0,\ \forall A \in S^*$ (vague statement), TP returns the uniform probability as P^* (that may be seen applying (b)).

Let us now suppose that $S \subsetneq S^*$, which is the *partial assessment* case.

Most known transformations cannot be *directly* applied to partial assessments, because they are based on quantities which are typically defined on the whole 2^{Ω}. They might anyway be applied *indirectly*, extending the coherent lower probability to 2^{Ω}. As discussed in Section 4.3, the natural extension $\underline{P}_E$ appears to be the most appropriate extension both theoretically and computationally. A transformation requiring a complete assignment could then be applied to $\underline{P}_E$ on 2^{Ω}, with the same limitations discussed, for each case, in Section 5.1.

The transformation we are proposing may be applied *directly* to partial assessments, as far as both lower and upper probability values are assigned for each $A \in S$.

If this condition holds, it suffices to replace S^* with S in TP; otherwise the natural extension (of either $\underline{P}$ or $\overline{P}$) on $S' = S \cup \{A : A^c \in S\}$ should be computed before replacing S^* with S'. In both cases the transformation problem still returns a unique coherent precise probability $P^*(A)$, $\forall A$ ($A \in S$ or $A \in S'$, respectively). Note that $P^*(\omega_1), \ldots, P^*(\omega_N)$ are generally not uniquely determined in the partial assessment case (except for the atoms included in S or in S').

We give now an example of the difficulty pointed out in Section 4.3. In particular, we consider a case where the coherent extension to some event(s) in S^* is unique and show the consequences of ignoring this piece of information.

Example. Given $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and $S = \{\omega_1, \omega_1 \cup \omega_2\}$, the assignment $\underline{P}(\omega_1) = \underline{P}(\omega_1 \cup \omega_2) = a$, $\overline{P}(\omega_1) = \overline{P}(\omega_1 \cup \omega_2) = 1 - a$, ($a \in [0, \frac{1}{3}]$) on S is equivalent to the lower probability assignment $\underline{P}(\omega_1) = \underline{P}(\omega_1 \cup \omega_2) = \underline{P}(\omega_3) = \underline{P}(\omega_2 \cup \omega_3) = a$ on the set $S^L = \{\omega_1, \omega_1 \cup \omega_2, \omega_3, \omega_2 \cup \omega_3\}$, which is easily seen to be coherent (for instance, using the envelope theorem). Since upper and lower probabilities are given for every event in S, we may consider finding P^* in a direct way. Here

[4] Non-emptiness is implied by coherence, which ensures that $\mathscr{M} \neq \emptyset$ in the lower envelope theorem.

$P_m(\omega_1) = P_m(\omega_1 \cup \omega_2) = \frac{1}{2}$ is a coherent probability on S (being the restriction on S of a probability on 2^Ω obtained from $P_m(\omega_1) = P_m(\omega_3) = \frac{1}{2}$, $P_m(\omega_2) = 0$), hence $P^* = P_m$.

However, using (3) to obtain $\underline{P}(\omega_1 \cup \omega_2) \geqslant \underline{P}(\omega_1) + \underline{P}(\omega_2)$ and since $\underline{P}(\omega_1 \cup \omega_2) = \underline{P}(\omega_1) = a$, we note that the given $\underline{P}$ has a unique coherent extension on ω_2, $\underline{P}(\omega_2) = \underline{P}_E(\omega_2) = \underline{P}_U(\omega_2) = 0$. Since $\underline{P}(\omega_2)$ is determined by the assessment on S, we consider computing P^* starting from $S^+ = S \cup \{\omega_2\}$. We therefore add $\underline{P}(\omega_2) = 0$ and, to be able to apply the transformation to the new assignment, $\overline{P}_E(\omega_2) = 1 - 2a$ to the initial assessment (note that the initial assignment does not entail a unique $\overline{P}(\omega_2)$). The new P_m is no longer a coherent probability ($P_m(\omega_1) = P_m(\omega_1 \cup \omega_2) = \frac{1}{2}$, $P_m(\omega_2) = \frac{1}{2} - a$, hence P_m is not additive), and we may compute P^* noting that the global minimum of $\varphi = (P^*(\omega_1) - \frac{1}{2})^2 + (P^*(\omega_2) - \frac{1}{2} + a)^2 + (P^*(\omega_1) + P^*(\omega_2) - \frac{1}{2})^2$ satisfies (12), (14) and therefore (by property (b) of TP) gives the required P^*, which is such that $P^*(\omega_1) = \frac{1+a}{3}$, $P^*(\omega_2) = \frac{1-2a}{3}$. Summing up, we obtain:

$$P^*(\omega_1) = P^*(\omega_1 \cup \omega_2) = \frac{1}{2}, \text{ operating on } S$$

$$P^*(\omega_1) = \frac{1+a}{3}, \quad P^*(\omega_1 \cup \omega_2) = \frac{2-a}{3}, \text{ operating on } S^+$$

To get an idea of the difference, let $a = 0$. $\underline{P}$ is then vague, and its most intuitive transformation appears to be (the restriction on S of) the uniform probability P_{unif}. However P^* is equal to P_{unif} when working on S^+, not when using S.

Clearly, the example above is not sufficient to infer what implications of a given assessment should be necessarily considered before running the transformation. For instance, it is not even simple in general to detect a priori (i.e. without computing upper and lower extensions) those events, if any, which allow a unique extension of $\underline{P}$, and this task may be not necessarily simpler than just computing $\underline{P}_E$ for all $A \notin S,\ A \in S^*$.

6. Transformations into possibilities

6.1. Previous proposals

As to our knowledge, the problem of transforming an imprecise probability into a possibility has not been considered in general.

Transformations of probability intervals given on atoms into a possibility distribution are considered in [13]. This is a restricted case of partial probability assignment and relies in particular on the assumption that a complete partition is initially given. Three procedures are defined by suitably extending previous proposals of transformations from precise probabilities into possibilities, they cannot however be applied to generic partial assignments.

The special case of approximating a belief function with a possibility has been thoroughly analyzed in [20], considering both inner and outer consonant approximations. Inner approximations have been dealt with subsequently in [31,32]. An outer consonant approximation of a belief function *Bel* is a necessity measure N such that $N(A) \leqslant Bel(A)$ (equivalently, $\Pi(A) \geqslant Pl(A)$), $\forall A \in 2^{\Omega}$, while reversing the inequalities we obtain an inner consonant approximation.

We focus on outer approximations, as they satisfy the consistency criterion. In [20] a heuristic algorithm is provided to derive an outer consonant approximation obeying certain minimality requirements. The extension of the results of [20] to the case of capacities is discussed in [3], this issue is however beyond the scope of the present paper. In fact, the transformations considered in [20,3] are based on the criteria of specificity preservation and quantitative similarity introduced above, while neglecting preference preservation. It has however to be noted that preference preservation might be considered the most important criterion for this kind of information interchange, due to the basically ordinal nature of possibility measures. As stated in [18]: "Possibility and necessity measures are set-functions that can provide simple ordinal representations of graded belief. Their particular character lies in their ordinal nature". Hence, it is reasonable to assume that a software agent adopting possibility theory as its uncertainty model is mainly interested in the credibility ordering associated with other agents' evaluations, and would possibly accept relatively loose approximations of the sender's numerical evaluation. A transformation procedure from imprecise probabilities into possibilities based on preference preservation for an arbitrary set of interesting events was introduced by the authors in [5]. In [28] a somewhat similar criterion, concerning the preservation of fuzzy T-preorders on elementary events only, was considered for the special case of approximating complete belief functions. In this paper we propose an improved and simpler version of the procedure in [5], showing that it produces a possibility with minimal additional imprecision among those satisfying preference preservation. Since the procedure works indifferently on partial and complete possibilities, we first consider the issue of characterizing partial possibilities.

6.2. Partial possibilities

As well known (e.g. see [19]), a possibility measure Π on the powerset 2^{Ω} of a finite partition $\Omega = \{\omega_1, \ldots, \omega_N\}$ can be defined by assigning a possibility distribution $\pi(\cdot) : \Omega \to [0, 1]$. Then $\Pi(\emptyset) = 0$ and, $\forall A \in 2^{\Omega} \setminus \{\emptyset\}$,

$$\Pi(A) = \max_{\omega_i \in A} \pi(\omega_i) \tag{15}$$

We will consider only *normal* distributions, i.e. such that $\exists \omega_j \in \Omega : \pi(\omega_j) = 1$.

As to our knowledge, the notion of a partial possibility measure defined on an arbitrary set of events included in 2^{Ω} has first been considered in [50]. Further generalizations, concerning partial possibility measures whose codomain is a lattice [14,35], are beyond the scope of the present paper.

The definition in [50] assumes that a reference partition Ω is given a priori. In [5] we slightly generalized this definition: in fact an arbitrary set of events[5] $S = \{A_1, \ldots, A_n\}$ may or may not be defined starting from an underlying partition Ω, but in any case the *partition* Ω_G *generated* by $A_1, \ldots, A_n$ can be obtained as the set of all intersections $A'_1 \cap \cdots \cap A'_n$, where each A'_i is alternatively replaced by either A_i or its complement A_i^c. Those intersections that are not empty constitute the atoms of Ω_G. Partition Ω_G satisfies the following properties:

(i) any event $A_i \in S$ is a union of some atoms of the partition (those included into A_i);
(ii) Ω_G is the coarsest partition with the property (i).

We recall that a partition Ω is *coarser* than Ω' (or, equivalently, Ω' is *more refined* than Ω) iff every atom of Ω is a union of atoms of Ω'.

Property (i) is useful to relate partial to ordinary possibilities, as is done in the following definition.

Definition 6. $\Lambda : S \to [0,1]$ is a partial possibility (on S) iff there exists a possibility measure $\Pi : 2^\Omega \to [0,1]$, such that $\forall A \in S,\ \Lambda(A) = \Pi(A)$, where 2^Ω is the powerset of a finite[6] partition Ω satisfying property (i) above.

The next lemma ensures that the concept of partial possibility we propose is well-defined, in the sense that it does not depend on the choice of Ω within the class of the finite partitions having the property (i). In particular, one may refer to Ω_G, i.e. the coarsest of these partitions.

Lemma 7. $\Lambda : S \to [0,1]$ *is a partial possibility with respect to a possibility defined on* 2^{Ω_G} *iff it is so with respect to a possibility defined on* 2^Ω, *where* Ω *is a finite partition more refined than* Ω_G.

Proof. We show that whenever Λ is the restriction on S of a possibility measure Π on 2^Ω, then it is also the restriction of a possibility measure Π' on 2^{Ω_G}. In fact, let Λ be a partial possibility with respect to Π, and let π be the underlying possibility distribution defined on $\Omega = \{\omega_1, \ldots, \omega_t\}$. Any event of S is a union of atoms of $\Omega_G = \{\omega_1^G, \ldots, \omega_m^G\}$: consider one of them, let it be $A_i = \omega_1^G \cup \cdots \cup \omega_k^G,\ k < m$. Then we have

$$\Pi(A_i) = \max_{\omega_h \in A_i} \pi(\omega_h) = \max\left(\max_{\omega_h \in \omega_1^G} \pi(\omega_h), \ldots, \max_{\omega_h \in \omega_k^G} \pi(\omega_h)\right).$$

Now put $\pi'(\omega_i^G) = \max_{\omega_h \in \omega_i^G} \pi(\omega_h)$ to assign possibility values to the atoms $\omega_1^G, \ldots, \omega_k^G$ of Ω_G. By considering all the events in S, this assignment can be easily

[5] Again we assume that $\emptyset$ and Ω do not belong to S.

[6] The finiteness assumption can be easily dropped, also in the following Lemma 7. However, only finite partitions are of interest in this paper.

extended to other atoms to obtain a complete possibility distribution π' on Ω_G, and, applying (15), a possibility measure Π' on 2^{Ω_G}. Clearly, by construction we have $\Lambda(A_i) = \Pi'(A_i), \forall A_i \in S$, i.e. Λ is a partial possibility with respect to Ω_G.

Conversely, if Λ is a partial possibility, the reference partition is Ω_G, π' is the corresponding possibility distribution, and Ω is any partition more refined than Ω_G, to see that Λ is a partial possibility also referring to Ω it is sufficient to obtain Λ from an appropriate possibility distribution $\pi : \Omega \to [0,1]$ using (15). This can be done in more ways; a simple choice is to put, for every $\omega_i^G \in \Omega_G$, $\pi(\omega_j) = \pi'(\omega_i^G)$, $\forall \omega_j \in \Omega : \omega_j \in \omega_i^G$. $\square$

It is shown in [50] that *P-consistency* is a necessary and sufficient condition for a partial assessment Λ to be a partial possibility:

Definition 8. $\Lambda : S \to [0,1]$ is P-consistent iff for any $A \in S$, and for any family $(A_k)_{k\in K}$ of elements of S:

$$A \subseteq \bigcup_{k\in K} A_k \Rightarrow \Lambda(A) \leqslant \sup_{k\in K} \Lambda(A_k) \tag{16}$$

In [5] we provided a characterization of partial possibilities giving a simpler answer to the problem of establishing whether a given uncertainty assessment Λ on S is a partial possibility. In fact, while Definition 8 refers to all families of elements of S, the equivalent conditions in the following Proposition 9 involve some specific families only, after ordering the values of Λ.

Suppose, for ease of notation, that $0 \leqslant \Lambda(A_1) \leqslant \cdots \leqslant \Lambda(A_n) \leqslant 1$.

Further let $v_1 < \cdots < v_m, m \geqslant 1$, be the distinct values assumed by $\Lambda(\cdot)$, and define for $j = 1, \ldots, m$:

$$V_j^* = \cup(A_i : \Lambda(A_i) = v_j) \tag{17}$$

$$V_j = \bigcup_{k=1}^{j} V_k^*; \quad V_0 = \emptyset \tag{18}$$

$$SV_j^* = \{A_i : \Lambda(A_i) = v_j\} \tag{19}$$

Proposition 9. *An uncertainty assessment $\Lambda : S \to [0,1]$ is a partial possibility iff both of the following conditions hold*:

$$\Lambda(A_n) = 1 \vee V_m \neq \Omega \tag{20}$$

$$\forall j,\ \forall A_i \in SV_j^*, \quad \exists \omega_k^G \in \Omega_G : \omega_k^G \in A_i \cap V_{j-1}^c \tag{21}$$

Proof. Suppose first that both (20) and (21) hold. A possibility distribution function $\pi(\cdot)$ can be assigned on Ω_G as follows:

(1) for $j = 1, \ldots, m$, for every $A_i \in SV_j^*$ select one $\omega_k^G \in \Omega_G : \omega_k^G \in A_i \cap V_{j-1}^c$ and put $\pi(\omega_k^G) = \Lambda(A_i)$;
(2) if $\Lambda(A_n) < 1$, by (20) it is $V_m = A_1 \cup \cdots \cup A_n \neq \Omega$, and we have $\omega_0^G = A_1^c \cap \cdots \cap A_n^c \neq \emptyset$, $\omega_0^G \in \Omega_G$. We put then $\pi(\omega_0^G) = 1$;
(3) assign $\pi(\cdot) = 0$ to all the remaining atoms of Ω_G to complete the definition of the possibility distribution π on Ω_G.

Then the possibility measure Π obtained from π applying (15) is such that $\Pi(A_i) = \Lambda(A_i), \forall A_i \in S$.

Conversely, suppose now that either (20) or (21) do not hold.

If (20) is false, it is $\Lambda(A_n) < 1 \wedge V_m = \Omega$. Hence every atom of Ω_G is included into (at least) one $A_i \in S$, and whatever possibility distribution π is assigned on Ω_G, this holds in particular for every ω_j^G satisfying the normality condition $\pi(\omega_j^G) = 1$. Therefore, by (15) at least one $A_i \in S$ should be given $\Pi(A_i) = 1$ thus yielding $\Pi(A_i) \neq \Lambda(A_i)$.

If (21) does not hold, this means that $\exists A_i \in SV_j^*,\ j > 1$ such that $A_i \cap V_{j-1}^c = \emptyset$, i.e. $A_i \subseteq V_{j-1}$. Then Λ is not a partial possibility being not P-consistent: (16) does not hold with $A = A_i$ and (A_k) formed by those and only those $A_k \in S$ such that $\Lambda(A_k) \leqslant v_{j-1}$. $\square$

Note that condition (21), involving atoms of the generated partition, can be checked exploiting only information about implication and incompatibility among events, by considering its negation (as in the proof above), namely:

$$\exists j,\ \exists A_i :\ \Lambda(A_i) = v_j, \quad A_i \subseteq V_{j-1} \tag{22}$$

Using (22) we may conclude that (21) may be operationally checked by verifying whether the set $NOTSAT$ is empty or not, where

$$NOTSAT(\Lambda) = \{A_i | \Lambda(A_i) = v_j,\ A_i \subseteq V_{j-1}\} \tag{23}$$

6.3. A transformation procedure based on preference preservation

Given a coherent upper probability[7] $\overline{P}$ on an arbitrary set $S = \{A_1, \ldots, A_n\}$, we define a procedure producing a partial possibility Π such that:

$$\forall A_i \in S, \quad \Pi(A_i) \geqslant \overline{P}(A_i) \tag{24}$$

$$\forall A_i,\ A_j \in S, \quad \overline{P}(A_i) = \overline{P}(A_j) \Rightarrow \Pi(A_i) = \Pi(A_j) \tag{25}$$

$$\overline{P}(A_i) > \overline{P}(A_j) \Rightarrow \Pi(A_i) \geqslant \Pi(A_j) \tag{26}$$

[7] Actually the procedure is applicable to any capacity, since its properties only depend on monotonicity of $\overline{P}$.

Condition (24) enforces the consistency principle (Section 4.1), while (25) and (26) impose a weak form of the preference preservation principle. The procedure we propose is based on Proposition 9, namely on conditions (20) and (21). All the implication and incompatibility relations among the events of S are assumed to be known, as well as whether $\cup_{A_i \in S} A_i = \Omega$.

Checking whether condition (20) is satisfied is straightforward. If not, it can easily be accommodated, by putting $\Pi(A_i) = 1,\ \forall A_i : \overline{P}(A_i) = v_m$; this also complies with (25). The modified assignment has to be processed further if condition (21) is not satisfied. As to this point, the transformation algorithm should modify $\overline{P}$ to make the set $NOTSAT$ defined in (23) empty.

Suppose first that $NOTSAT$ includes just one event A_i. In this case, the value assigned to at least one event B included in V_{j-1} should be raised to v_j so that $A_i \nsubseteq V_{j-1} \backslash B$, and, to ensure preference preservation, one should also raise to v_j the values of all events C such that $\overline{P}(B) \leqslant \overline{P}(C) < v_j$. Of course, to minimize the additional imprecision, the initial value of $\overline{P}(B)$ should be as high as possible.

If $NOTSAT$ includes more than one event, the matter is more complicated. Several modifications of this kind should be enforced, but it is not immediate to devise which is the "right" set of modifications, as they cannot be considered independently: each modification may affect several of the sets V_i and, in a sense, the modifications interact with one another.

To start reasoning, let us note that if a given V_i includes an event $A \in SV_j^*$, with $j > i$, then all (due to preference preservation) events included in SV_i^* should at least be raised to the value v_j. Then, in turn, if V_j includes an event $B \in SV_k^*$ with $k > j$, all events in SV_j^*, along with all those previously raised to the value v_j, should be raised to v_k, and so on.

As a simple example, consider $\overline{P}(\omega_1 \cup \omega_2) = \alpha > \overline{P}(\omega_3 \cup \omega_4) = \beta > \overline{P}(\omega_1) = \overline{P}(\omega_2) = \gamma > \overline{P}(\omega_3) = \overline{P}(\omega_4) = \delta$. Starting with the lowest values, the one assigned to both ω_3 and ω_4 should be raised from δ to β to satisfy the basic properties of possibility measures. Also the value assigned to both ω_1 and ω_2 should be raised from γ to β due to preference preservation. However, the value assigned to both ω_1 and ω_2 should be definitely raised to α and, due to preference preservation, this entails that also the value of all other events considered in the example should in turn be raised to α.

When proceeding sequentially, starting from the events in $NOTSAT$ with lowest values for $\overline{P}$, the chain of interacting modifications involves sets V_j with progressively higher indexes j. The chain stops when a V_{stop} is first met such that $\nexists A | A \subseteq V_{stop} \wedge A \in SV_q^*,\ q > stop$. Such a V_{stop} acts as a milestone for the set of related modifications: other modifications might possibly be necessary, but if so, they will involve *only* events with higher $\overline{P}$ value than those in V_{stop}.

This idea gives rise to the following algorithm.

Step 1.
$\forall A_i \in S$ `put`: $\Pi(E) := \overline{P}(E)$;
Step 2.
`if` $(\cup(A_i : A_i \in S) = \Omega \wedge \max_{A_i \in S} \Pi(A_i) \neq 1)$

`then` $\forall A_i : \Pi(A_i) = \max_{B \in S} \Pi(B)$, `put`: $\Pi(A_i) := 1$;
Step 3.
`put`: *index*: = 0;
BEGIN MAIN LOOP
Step 4.
`put`: $v_1 < \cdots < v_m$ the distinct values assumed by Π on S
`put`: $V_j^* := \cup(A_i : \Pi(A_i) = v_j); V_j := \cup_{1 \leqslant k \leqslant j} V_k^*$
`for` $j = 1, \ldots, m$, `put`: $TOLIFT := \{j > index | \exists A_i \in S, A_i \subseteq V_j, \Pi(A_i) > v_j\}$
`if` $TOLIFT = \emptyset$ `then` **EXIT**
`else`
Step 5.
`put`: $start := \min(TOLIFT); stop := \min\{j | j > start \wedge j \notin TOLIFT\}$
Step 6.
$\forall A_i | v_{start} \leqslant \Pi(A_i) < v_{stop}$, `put`: $\Pi(A_i) := v_{stop}$
Step 7.
`put`: *index*: = *stop*;   `goto` Step 4
END MAIN LOOP

Step 1 initializes the values of Π, Step 2 checks condition (20) and enforces it if needed, Step 3 initializes the variable *index* used in the main loop. Step 4 updates a set of variables on the basis of the current values of Π and *index*. In particular, if the set *TOLIFT* is not empty, the current Π is not a partial possibility. In fact, each index i included in *TOLIFT* corresponds to a set V_i for which condition (22) holds. If *TOLIFT* is empty, the procedure terminates, otherwise the following steps are executed.

Step 5 assigns two variables: the *start* index selects, among the sets V_i identified by *TOLIFT*, the one corresponding to the smallest value of the current Π, while the *stop* index identifies the minimal (with respect to set inclusion) of the sets V_j, termed V_{stop}, for which condition (22) does not hold.

The *stop* index is always well defined: at least $m \notin TOLIFT$, given that $\nexists A \in S | \Pi(A) > v_m$.

Step 6 increases to v_{stop} the possibility of events whose previous Π values are between v_{start} and v_{stop}. This corresponds to a set of related modifications as explained above. Since the possibility values up to v_{stop} will not need to be modified further, in Step 7 the variable *index* is increased to *stop* so that only higher values are considered in subsequent steps. Then the procedure goes back to the beginning of the main loop.

Proposition 10. *The above described algorithm is such that*:

(a) *it terminates*;
(b) *it returns a partial possibility* Π;
(c) *let* Π' *be any partial possibility consistent with* $\overline{P}$ *and respecting preference preservation* (*i.e. conditions* (24)–(26) *hold*, *where* Π *is replaced by* Π'). *Then* $\Pi(A) \leqslant \Pi'(A), \ \forall A \in S$.

Proof. To prove (a), suppose that $i \in TOLIFT$ at iteration (k) of the main loop. Letting $v_h = \max_{A \subseteq V_i, \Pi(A) > v_i} \Pi(A)$, B such that $\Pi(B) = v_h$, it holds that $\forall j | i \leqslant j < h,\ j \in TOLIFT$. In fact, $\Pi(B) > v_j$ and $B \subseteq V_i \subset V_j$. Then $stop \geqslant h$. This implies that any i such that $start \leqslant i < stop$ at iteration (k) does no longer belong to $TOLIFT$ at iteration (k + 1). On the other hand, there is no index in $TOLIFT$ at iteration (k + 1) which was not already in $TOLIFT$ at iteration (k). In other words, at least the index *start* is subtracted from $TOLIFT$ at each iteration, while no index is ever added: therefore the cardinality of $TOLIFT$ decreases at each iteration and algorithm termination is guaranteed.

To prove (b), it is sufficient to verify that $TOLIFT = \emptyset \iff NOTSAT = \emptyset$. In fact, $TOLIFT = \emptyset$ means that if $\Pi(A) = v_j > v_i$ then $A \nsubseteq V_i,\ \forall i < j$. In particular $A \nsubseteq V_{j-1}$ and therefore $NOTSAT = \emptyset$. On the other hand $NOTSAT = \emptyset$ means that if $\Pi(A) = v_j$, then $A \nsubseteq V_{j-1}$. Since $V_i \subseteq V_{j-1}$ for any $i < j$, then $A \nsubseteq V_i$, for any $i < j$. Therefore $TOLIFT = \emptyset$.

To prove (c), note that condition $\Pi(A) \leqslant \Pi'(A)$ obviously needs to be verified only for the events A such that $\Pi(A) > \overline{P}(A)$, namely such that $v_{start} \leqslant \overline{P}(A) < v_{stop}$ in one of the iterations (hence, by Step 6 of the algorithm, $\Pi(A) = v_{stop}$). Since $stop \notin TOLIFT$, it holds that $v^k_{stop} < v^{k+1}_{start}$ where v^k denotes the value assumed by v at iteration (k) of the algorithm. Therefore, the sets of events involved by each iteration are disjoint and it is sufficient to show that the desired condition holds considering a generic iteration of the algorithm (but for the last one, where $TOLIFT$ is already empty).

To this purpose, let A_{start} be an event such that $\overline{P}(A_{start}) = v_{start}$, namely an event with minimum initial value among those involved at iteration (k). It is then sufficient to show that $\Pi'(A_{start}) \geqslant v_{stop}$, since preference preservation entails then $\Pi'(A) \geqslant v_{stop}$ for all other events A involved by the iteration.

Let us note first that $\exists B_0$ which belongs to $NOTSAT$ at the iteration (k) we are considering and $B_0 \subseteq \cup_{\overline{P}(C) \leqslant v_{start}} C \wedge \overline{P}(B_0) > v_{start}$. When replacing Λ with the partial possibility Π' in (23), it must be $NOTSAT(\Pi') = \emptyset$ and, in particular, $B_0 \notin NOTSAT$. Therefore there exists $D_0 \in S$ such that

$$\overline{P}(D_0) \leqslant v_{start}, \quad \Pi'(D_0) \geqslant \Pi'(B_0) \geqslant \overline{P}(B_0) \tag{27}$$

Preference preservation and (27) imply

$$\Pi'(A_{start}) \geqslant \Pi'(D_0) \geqslant \overline{P}(B_0) \tag{28}$$

In the case $\overline{P}(B_0) = v_{stop}$, this is already the desired result. Otherwise, assuming $\overline{P}(B_0) = v_h < v_{stop}$, we have that $h \in TOLIFT$. Therefore $\exists B_1$ which belongs to $NOTSAT$ at iteration (k), $B_1 \subseteq \cup_{\overline{P}(C) \leqslant \overline{P}(B_0)} C \wedge \overline{P}(B_1) > \overline{P}(B_0)$. Reasoning as above, there exists $D_1 \in S$ such that $\overline{P}(D_1) \leqslant \overline{P}(B_0), \Pi'(D_1) \geqslant \Pi'(B_1) \geqslant \overline{P}(B_1)$. These inequalities, preference preservation, (27) and (28) imply $\Pi'(A_{start}) \geqslant \Pi'(D_0) \geqslant \Pi'(B_0) \geqslant \Pi'(D_1) \geqslant \Pi'(B_1) \geqslant \overline{P}(B_1)$. Again, either $\overline{P}(B_1) = v_{stop}$ or the argument can be iterated until a $\overline{P}(B_q) = v_{stop}$ is reached, yielding $\Pi'(A_{start}) \geqslant \Pi'(B_0) \geqslant \Pi'(B_1) \geqslant \cdots \geqslant \Pi'(B_q) \geqslant \overline{P}(B_q) = v_{stop}$. $\square$

The preceding proposition ensures that the algorithm outputs the best (in the sense of as precise as possible) partial possibility among those complying with the consistency and preference preservation principles (24)–(26). It can also be shown (we omit the lengthy proof) that the algorithm is equivalent to the one we introduced in [5], which is less immediate and generally requires a higher number of loops to terminate.

7. Transformations into belief functions

The problem of transforming a lower probability assessment $\underline{P}$ on 2^{Ω} into a belief function *Bel* was considered in [4], applying the *consistency* and *quantitative similarity* criteria. The consistency criterion is used regarding belief functions as less precise than imprecise probabilities. That can be justified by the following inferential argument, showing that belief functions may produce less precise inferences: if $\underline{P}$ is an unconditional lower probability and $\underline{P}(B) > 0$, it is known [47] that its vaguest (or least-committal) extension on $A|B$ is such that $\underline{P}(A|B) \geqslant \underline{P}(A \cap B)/(\underline{P}(A \cap B) + \overline{P}(A \cap B^c))$, with equality holding if $\underline{P}$ is a belief function (actually, it suffices that $\underline{P}$ is 2-monotone).

The imprecise probability/belief function transformation problem has been discussed also in [27], requiring the same consistency and quantitative similarity criteria as above. In particular, quantitative similarity is applied by minimizing the same objective function as in [4]. A computationally simpler heuristic algorithm, Iterative Rescaling Algorithm or IRM, is further discussed; however, as shown in [27], IRM guarantees consistency but does not minimize the objective function, in general.

In this section we consider the case of partial belief functions, which is not encompassed in the above mentioned papers. A partial belief on $S \subset 2^{\Omega}$ (if no partition Ω underlying S is given, the generated partition Ω_G can be considered) is an uncertainty assessment which is the restriction on S of some belief function defined on 2^{Ω}.

Partially specified belief functions have been considered in [36,39]. The work in [36] regards the special case of a partial assignment on the atoms of a partition and identifies necessary and sufficient conditions such that the assignment can be extended to a complete belief function. In [39] the more general problem of extending a partial assignment on an arbitrary set of events S to a belief function with minimum specificity is considered. A linear programming formulation of this problem is provided: it is remarked that a solution does not always exist and may not be unique, when it does. Simplified methods are then provided for some special cases, especially when the focal sets of the complete belief functions are a subset of S. Therefore, apart from special situations, even verifying whether a partial assignment can be extended to a belief function involves linear programming.

We turn now to the problem of transforming a coherent lower probability on S into a partial belief function. This is achieved by the following linear programming problem, that enforces quantitative similarity by minimizing function φ, which is linear in the variables $m_B(A),\ A \in 2^{\Omega}$. The problem constraints are due to consistency and to the properties of the mass function m_B of *Bel* (cf. also (1) and Proposition

1(c)). The transformation of an assignment on the whole 2^Ω is recovered from this problem putting $S = 2^\Omega \setminus \{\emptyset, \Omega\}$ and corresponds to that considered in [4].

$$\min \varphi = \sum_{E \in S} \left(\underline{P}(E) - \sum_{A \subseteq E} m_B(A) \right) \tag{29}$$

subject to:

$$\sum_{A \subseteq E} m_B(A) \leqslant \underline{P}(E), \ \forall E \in S; \quad m_B(A) \geqslant 0, \ \forall A \in 2^\Omega; \quad \sum_{A \in 2^\Omega} m(A) = 1 \tag{30}$$

The feasible region of this problem is always non-empty ($m_B(E) = 0$, $\forall E \neq \Omega, m_B(\Omega) = 1$ is a feasible point), but the solution is not necessarily unique (indeed, if there is more than one solution there are infinitely many, since the feasible region is convex). However, we note that since any solution minimizes φ, by (29) it also maximizes $\sum_{E \in S} \sum_{A \subseteq E} m_B(A) = \sum_{E \in S} Bel(E)$. Therefore no one of the solutions for problem (29), (30) dominates or is dominated by any other solution on the events in S: there is no reason for preferring one solution rather than another one on dominance grounds. As a possible additional selection criterion, minimum specificity [39] could be adopted; this requires however solving an additional linear programming problem.

As well known, given a reference partition with cardinality N, problems with belief functions may involve a considerably higher number of variables than problems with precise probabilities or possibilities, which are both determined by at most N distinct values. An approach aiming to reduce the number of variables in (29), (30) might be to search for a belief function Bel^* such that $m_B^*(A) = 0$ whenever $|A| > k$ $(1 < k < N = |\Omega|)$ and $A \neq \Omega$. This approach is closely related to idea of k-additive belief function [25]: Bel^* may not be k-additive only because $m_B^*(\Omega)$ is not necessarily 0; we call Bel^* *quasi* k-additive. It simplifies the problem (29), (30) while keeping its feasible region non-empty (since it contains at least the solution $m_B^*(E) = 0, \ \forall E \neq \Omega, \ m_B^*(\Omega) = 1$).

Note also that the solution of (29), (30) in terms of either $m_B(\cdot)$ or $m_B^*(\cdot)$ automatically generates a belief function on the whole 2^Ω and it is left to the final user (the receiving agent) whether to use only its restriction on S or not.

8. An example

Some gamblers consider betting on $S = \{E_1, \ldots, E_{10}\}$, where the events in S concern a certain future football World Cup and are now described. We call *titled* (*untitled*) a team which already (never) won the Cup in the past.

- E_1: the winner is a titled European team;
- E_2: the winner is an untitled European team;
- E_3: the winner is an untitled South American team;
- E_4: the winner is a titled team;

- E_5: the winner is either an untitled European team or a titled South American team;
- E_6: the winner is a European team;
- E_7: the winner is either a titled South American team or a European team;
- E_8: the winner is an African team;
- E_9: the winner is a team from an Arab-speaking country;
- E_{10}: the winner is a team from Africa or from an Arab-speaking country, and took part in some past edition of the Cup.

The partition generated by these events is formed by the atoms $e_1, \ldots, e_{11}$ shown in Fig. 1. Note that $E_1 = e_1,\ E_2 = e_2,\ E_3 = e_3,\ E_4 = e_1 \cup e_4,\ E_5 = e_2 \cup e_4,\ E_6 = e_1 \cup e_2,\ E_7 = e_1 \cup e_2 \cup e_4,\ E_8 = e_5 \cup e_6 \cup e_7 \cup e_8,\ E_9 = e_7 \cup e_8 \cup e_9 \cup e_{10},\ E_{10} = e_6 \cup e_7 \cup e_9$, and that $e_{11} = E_1^c \cap E_2^c \cap \cdots \cap E_{10}^c$.

An influential agency publishes on its web site the assignment of upper and lower coherent probabilities on S shown in Table 1.

Some of the gamblers adopt precise probabilities, possibilities or belief functions as their own uncertainty representations. To make use of the agency evaluations they need transforming them into their representations.

We do that, exemplifying the procedures proposed in the previous sections. Linear and quadratic programming computations were carried out using Matlab on a standard PC.

As for transforming the given assessment into a precise probability, problem TP (13), (14) gives the following result: $P^*(e_1) = 0.2115$; $P^*(e_2) = 0.094$; $P^*(e_3) = 0.075$; $P^*(e_4) = 0.228$; $P^*(e_5) = 0.035$; $P^*(e_6) = 0$; $P^*(e_7) = 0.19$; $P^*(e_8) = 0.005$; $P^*(e_9) = 0$; $P^*(e_{10}) = 0$; $P^*(e_{11}) = 0.1615$.

Referring to the events in S, the values of P^* are shown in Table 2.

A straightforward computation shows that $P^* \neq P_m$ on S and that, in this case, the difference is relatively small ($\max_{A \in S}|P_m(A) - P^*(A)| = 0.002$).

Considering now the transformation into possibilities, let us first apply the algorithm proposed in Section 6 to the values of $\overline{P}$ given above for S. First note that the union of the events in S is not equal to Ω, therefore Step 2 does not modify the assessment. There are eight distinct values ($v_1 = 0.1$; $v_2 = 0.11$; $v_3 = 0.24$; $v_4 =$

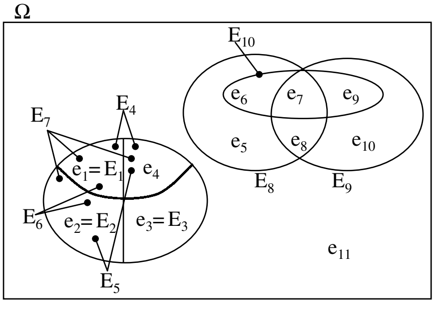

Fig. 1. Events and generated partition in the football example.

Table 1
A coherent imprecise probability assessment on S

	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	E_{10}
$\underline{P}$	0.18	0.08	0.05	0.43	0.28	0.29	0.52	0.18	0.15	0.06
$\overline{P}$	0.24	0.11	0.1	0.45	0.36	0.32	0.55	0.28	0.24	0.32

Table 2
Transformation result in the case of precise probability

	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	E_{10}
P^*	0.2115	0.094	0.075	0.4395	0.322	0.3055	0.5335	0.23	0.195	0.19

$0.28; v_5 = 0.32;\ v_6 = 0.36;\ v_7 = 0.45;\ v_8 = 0.55)$ in the initial assessment, the corresponding V_j^* are:

$$V_1^* = E_3; \quad V_2^* = E_2; \quad V_3^* = E_1 \cup E_9; \quad V_4^* = E_8;$$
$$V_5^* = E_6 \cup E_{10}; \quad V_6^* = E_5; \quad V_7^* = E_4; \quad V_8^* = E_7$$

The corresponding V_j are obtained by incremental union of V_j^*.

Then $TOLIFT = \{3,4,6,7\}$. In fact event E_6, having evaluation v_5, is included in both V_3 and V_4, while event E_7, with value v_8, is included in both V_6 and V_7. Accordingly, $start = 3,\ stop = 5$ at Step 5, and the values of E_1, E_9, and E_8 are raised to $v_5 = 0.32$. In the second main loop iteration, $TOLIFT = \{6,7\},\ start = 6,\ stop = 8$ and the values of E_5 and E_4 are raised to $v_8 = 0.55$. In the subsequent iteration $TOLIFT = \emptyset$ and the procedure terminates.

The resulting partial possibility is reported in Table 3.

As a remark, note that the same partial possibility would be obtained, for instance, for any initial evaluation such that $\overline{P}(E_5) \in]0.32, 0.45]$, while in the case $\overline{P}(E_5) = 0.32$ we would obtain $TOLIFT = \{3,4,5,6,7\}$ in the first iteration, and consequently all events but for E_2 and E_3 would get the value $v_8 = 0.55$. This is a further instance of the discontinuity of this transformation already pointed out in Section 4.2.

To fully exploit the information carried by the initial assessment, one should also take account of the lower probability assessments, which correspond (cf. (2)) to upper probabilities on the complementary events. This means that the transformation should be applied to the events in $S' = S \cup \{A | A^c \in S\}$. Doing so the restriction on S of the resulting possibility is less precise than the previously obtained possibility. In fact, the union of events in S' is now equal to Ω, therefore the value of one of

Table 3
Transformation result in the case of possibility

	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	E_{10}
Π	0.32	0.11	0.1	0.55	0.55	0.32	0.55	0.32	0.32	0.32

Table 4
Transformation results in the case of belief function

	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	E_{10}
Bel	0.18	0.08	0.05	0.4104	0.28	0.2896	0.52	0.18	0.15	0.06

the events is raised to 1. Then it can be seen that the events E_4, E_5, E_7 are given possibility value 1 in the main loop, while the other events in S get the same value as above.

Turning to the transformation into belief functions, solving problem (29), (30) using the lower probability assessments on the events of S gives the results shown in Table 4.

The initial assessment is not a partial belief function: the transformed assessment modifies the values of events E_4 and E_6. The overall difference between the initial and final evaluations is relatively small: the value of the objective function φ is 0.02. The size of this problem (2047 variables) still allowed to obtain a solution in a reasonable time (about 2 seconds) on a Pentium III (733 MHz) PC, mainly due to the small number of events in S. Significantly increasing the cardinality of S and/or of the generated partition would make the problem intractable. As mentioned in Section 7, a possible approach to limit the complexity of the problem consists in computing quasi-k-additive belief functions. In this example, using these functions with $k \geqslant 2$ kept on returning a value of the objective function of 0.02 and only slightly modified the values of the events E_4 and E_6. For instance, with $k = 2$ we get: $Bel^*(E_4) = 0.4218$, $Bel^*(E_6) = 0.2782$; with $k = 3$, $Bel^*(E_4) = 0.4160$, $Bel^*(E_6) = 0.2840$.

Also in this case, one may consider exploiting the whole initial assessment by taking account of the upper probabilities and solving the problem on S', whose cardinality is twice as much as the one of S. Computation times are still of the order of a few seconds and the restriction on S of the resulting belief function is very close to the one presented above: again $\varphi = 0.02$ and only the evaluations of E_4 (0.4240) and of E_6 (0.2760) are modified. Very similar results hold with quasi-k-belief functions.

Clearly, this and other conclusions should not be regarded as general achievements, since this example was primarily devised as an illustration of the proposed transformation procedures.

9. Discussion and conclusions

In this paper we have analyzed the issue of defining an interchange format for communication of information affected by uncertainty among heterogeneous software agents.

While interoperability among heterogeneous software agents is largely recognized as a key issue for the development of multi-agent systems, the problem of information interchange under uncertainty has received a relatively limited attention in the literature, most proposals being restricted to the case of communication of "certain" information. The only other approach to this problem we are aware of is presented in a series of papers [52,53,37], where the problem of defining transformation methods among different uncertainty representation approaches is considered. These works deal with the uncertainty models used in the EMYCIN, PROSPECTOR and MYCIN systems and do not consider more general theories. Moreover they do not introduce the notion of a common interchange format and therefore consider direct inter-formalism transformations, which is disadvantageous in some respects and in particular is not appropriate for indirect communication. Our work, while sharing the same basic motivations, extends the results of [52,53,37] by devising a suitable interchange format and by considering transformations involving more general formalisms for uncertainty representation.

We identified coherent imprecise probabilities as a suitable candidate for an uncertainty interchange format since they include as special cases several well-known formalisms, in particular precise probabilities, possibilities, belief functions, whose importance is largely recognized. Moreover, imprecise probabilities are defined on arbitrary set of events, a feature which turns out to be advantageous in the application context we consider.

We then investigated the problem of devising new transformation procedures from imprecise probabilities into the three formalisms mentioned above, as evidenced in the relevant sections, where comparisons with respect to the existing literature have also been carried out. In general, it turns out that existing procedures cover special cases of the transformations considered, but cannot be appropriately applied to the general case, nor do they handle the partial assessment case.

Further analyses of the proposed procedures, either at the theoretical or experimental level, are reserved for future work. In particular, it would be useful to characterize the information distortion introduced by the procedures, to carry out a deeper investigation about the trade-offs between the use of partial assignments or of their completion through natural extension, and to investigate the relationships between transformation procedures and inferential methods.

While the formalism we adopted appears to be sufficiently general to accommodate most application needs, we recall that, according to [49], other more general models, in particular *partial preference orderings* and *sets of desirable gambles*, should be considered in order to establish a unified theory of imprecise probability. Achievements in this area might provide suggestions for the definition of a more general interchange format and pose new problems as far as the issue of transformations is concerned.

A further direction of generalization concerns formalisms involving a higher order of imprecision with respect to the ones we consider. As mentioned in Section 3, a significant example, in the context of belief functions theory, is provided by the notion of imprecise belief structures [17]. In particular, the issue of computing a belief function with suitable properties from an imprecise belief structure is addressed in [17]: this kind of transformation involves, in a sense, a reduction of the order of imprecision within the same theory rather than a transformation between different formalisms. Therefore it concerns a distinct (and in a sense complementary) problem with respect to the one addressed in the present paper.

Finally, another interesting question concerns studying, in the context we consider, transformations based on the principle of uncertainty invariance [33], which however require the definition of a measure of uncertainty for imprecise probabilities: recent investigations in this area include [51,33,1,2,38].

References

[1] J. Abellan, S. Moral, Completing a total uncertainty measure in Dempster–Shafer theory, Int. J. Gen. Syst. 28 (1999) 299–314.
[2] J. Abellan, S. Moral, A non-specificity measure for convex sets of probability distribution, Int. J. Uncertain. Fuzz. Knowledge-Based Syst. 8 (2000) 357–367.
[3] P. Baroni, Extending consonant approximations to capacities, in: Proc. of IPMU 2004, Perugia, Italy, 2004, pp. 1127–1134.
[4] P. Baroni, P. Vicig, An uncertainty interchange format for multi-agent systems based on imprecise probabilities, in: Proc. of IPMU 2000, Madrid, Spain, 2000, pp. 1027–1034.
[5] P. Baroni, P. Vicig, Transforming imprecise probabilities into partial possibilities, in: Proc. of IPMU 2002, Annecy, France, 2002, pp. 1069–1076.
[6] P. Baroni, P. Vicig, On the conceptual status of belief functions with respect to coherent lower probabilities, in: Proc. of ECSQARU 2001, Toulouse, France, 2001, pp. 328–339.
[7] P. Baroni, P. Vicig, Transformations from imprecise to precise probabilities, in: Proc. of ECSQARU 2003, Aalborg, Denmark, 2003, pp. 37–49.
[8] M.S. Bazaraa, H.D. Sherali, C.M. Shetty, Nonlinear Programming, second ed., Wiley, New York, 1993.
[9] A. Chateauneuf, J.-Y. Jaffray, Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion, Math. Social Sci. 17 (1989) 263–283.
[10] J.J.-Y. Chen, S.W. Su, AgentGateway: a communication tool for multi-agent systems, Inform. Sci. 150 (3–4) (2003) 153–164.
[11] B.R. Cobb, P.P. Shenoy, A comparison of methods for transforming belief functions models, in: Proc. of ECSQARU 2003, Aalborg, Denmark, 2003, pp. 255–266.
[12] M. Daniel, Consistency of probabilistic transformations of belief functions, in: Proc. of IPMU 2004, Perugia, Italy, 2004, pp. 1135–1142.
[13] L.M. de Campos, J.F. Huete, Measurement of possibility distributions, Int. J. Gen. Syst. 30 (3) (2001) 309–346.
[14] G. de Cooman, Possibility theory I: The measure- and integral-theoretic groundwork, Int. J. Gen. Syst. 25 (1997) 291–323.
[15] B. de Finetti, Theory of Probability, vol. 1, Wiley, London, 1974.
[16] D. Denneberg, Non-additive Measure and Integral, Kluwer, Dordrecht, NL, 1994.
[17] T. Denœux, Reasoning with imprecise belief structures, Int. J. Approx. Reason. 20 (1999) 79–111.
[18] D. Dubois, H.T. Nguyen, H. Prade, Possibility theory, probability and fuzzy sets: misunderstandings, bridges and gaps, in: D. Dubois, H. Prade (Eds.), Fundamentals of Fuzzy Sets, Kluwer, Dordrecht, NL, 2000, pp. 343–438.
[19] D. Dubois, H. Prade, Possibility Theory, Plenum Press, New York, 1988.
[20] D. Dubois, H. Prade, Consonant approximations of belief functions, Int. J. Approx. Reason. 4 (1990) 419–449.
[21] D. Dubois, H. Prade, S. Sandri, On possibility/probability transformations, in: R. Lowen, M. Roubens (Eds.), Fuzzy Logic—State of the Art, Kluwer, Dordrecht, NL, 1993, pp. 103–112.
[22] T. Finin, Y. Labrou, J. Mayfield, KQML as an agent communication language, in: J. Bradshaw (Ed.), Software Agents, MIT Press, Boston, MA, 1997, pp. 291–316.
[23] FIPA 97 Specification, Part 2: Agent Communication Language. Available from: <http://www.fipa.org>.

[24] M.R. Genesereth, Knowledge Interchange Format. Draft proposed American National Standard (dpANS), NCITS.T2/98-004, 1998. Available from: <http://logic.stanford.edu/kif>.
[25] M. Grabisch, Upper Approximation of non-additive measures by k-additive measures—the case of belief functions, in: Proc. of ISIPTA'99, Ghent, Belgium, 1999, pp. 158–164.
[26] R.H. Guttman, A.G. Moukas, P. Maes, Agent-mediated electronic commerce: a survey, Knowl. Eng. Rev. 13 (2) (1998) 147–159.
[27] J.W. Hall, J. Lawry, Generation, combination and extension of random set approximations to coherent lower and upper probabilities, Reliab. Eng. Syst. Safety 85 (2004) 89–101.
[28] R. Hernandez, J. Recasens, On possibilistic and probabilistic approximations of unrestricted belief functions based on the concept of fuzzy T-preorder, Int. J. Uncertain. Fuzz. Knowledge-Based Syst. 10 (2002) 185–200.
[29] J.-Y. Jaffray, On the maximum-entropy probability which is consistent with a convex capacity, Int. J. Uncertain. Fuzz. Knowledge-Based Syst. 3 (1995) 27–33.
[30] N.R. Jennings, K. Sycara, M. Wooldridge, A roadmap of agent research and development, Autonom. Agents Multi-Agent Syst. 1 (1998) 7–38.
[31] C. Joslyn, G.J. Klir, Minimal information loss possibilistic approximations of random sets, in: Proc. of 1992 FUZZ-IEEE Conference, San Diego, CA, 1992, pp. 1081–1088.
[32] C. Joslyn, Possibilistic normalization of inconsistent random intervals, Advances in Systems Science and Applications, special issue, 1997, pp. 44–51.
[33] G.J. Klir, M.J. Wierman, Uncertainty-based Information. Elements of Generalized Information Theory, Physica-Verlag, Heidelberg, 1998.
[34] G.J. Klir, Uncertainty and information measures for imprecise probabilities: an overview, in: Proc. of ISIPTA'99, Ghent, Belgium, 1999, pp. 234–240.
[35] I. Kramosil, Partial lattice-valued possibilistic measures and some relations induced by them, in: Proc. of ECSQARU 2003, Aalborg, Denmark, 2003, pp. 395–406.
[36] J.F. Lemmer, H.E. Kyburg, Conditions for the existence of belief functions corresponding to intervals of beliefs, in: Proc. of AAAI 1991, Anaheim, CA, 1991, pp. 488–493.
[37] X. Luo, C. Zhang, H. Leung, Information sharing between heterogeneous uncertain reasoning models in a multi-agent environment: A case study, Int. J. Approx. Reason. 27 (2001) 27–59.
[38] J.-L. Marichal, Entropy of discrete Choquet capacities, Eur. J. Operat. Res. 137 (2002) 612–624.
[39] S. Moral, L.M. De Campos, Partially specified belief functions, in: Proc. of UAI 1993, Washington, DC, 1993, pp. 492–499.
[40] G.C. Rota, On the foundations of combinatorial theory. I. Theory of Möbius functions, Z. Wahrscheinlichkeit. Verwand. Gebiete 2 (1964) 340–368.
[41] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, NJ, 1976.
[42] P. Smets, Belief functions, in: P. Smets, E. Mamdani, D. Dubois, H. Prade (Eds.), Non-standard Logics for Automated Reasoning, Academic Press, London, 1988, pp. 253–277.
[43] P. Smets, Constructing the pignistic probability function in a context of uncertainty, in: M. Henrion, R.D. Shachter, L.N. Kanal, J.F. Lemmer (Eds.), Uncertainty in Artificial Intelligence 5, North Holland, Amsterdam, 1990, pp. 29–39.
[44] P. Smets, Decision making in the TBM: The necessity of the pignistic transformation, Int. J. Approx. Reason. 38 (2005) 133–147.
[45] P. Vicig, An algorithm for imprecise conditional probability assessments in expert systems, in: Proc. of IPMU 1996, Granada, Spain, 1996, pp. 61–66.
[46] F. Voorbraak, A computationally efficient approximation of Dempster–Shafer theory, Int. J. Man–Machine Studies 30 (1989) 525–536.
[47] P. Walley, Statistical Reasoning with Imprecise Probabilities, Chapman and Hall, London, 1991.
[48] P. Walley, Measures of uncertainty in expert systems, Artif. Intell. 83 (1996) 1–58.
[49] P. Walley, Towards a unified theory of imprecise probability, Int. J. Approx. Reason. 24 (2000) 125–148.
[50] Z. Wang, Extension of possibility measures defined on an arbitrary non-empty class of sets, in: Proc. of 1st IFSA Congress, Palma de Mallorca, Spain, 1985.

[51] R.R. Yager, On the entropy of fuzzy measures, IEEE Trans. Fuzzy Syst. 8 (4) (2000) 453–461.
[52] C. Zhang, Cooperation under uncertainty in distributed expert systems, Artif. Intell. 56 (1992) 21–69.
[53] C. Zhang, Heterogeneous transformation of uncertainties of propositions among inexact reasoning models, IEEE Trans. Knowl. Data Eng. 6 (3) (1994) 353–360.